RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: _10/520,401_

Source: _PCT/10_

Date Processed by STIC: _1/13/05_

ENTERED

PCT10

RAW SEQUENCE LISTING   DATE: 01/19/2005
PATENT APPLICATION: US/10/520,401   TIME: 20:32:42

Input Set : **N:\Crf4\01132005\J520401.raw**
Output Set: **N:\CRF4\01192005\J520401.raw**

```
 1 <110> APPLICANT: University degli Studi di Roma "La Sapienza"
 2       Consiglio Nazionale delle Ricerche
 3 <120> TITLE OF INVENTION: Purification, cloning and biochemical characterization of
XendoU,
 4       endoribonucleasic activity involved in Xenopus laevis small
 5       nuclear RNA splicing independent biosynthesis
 6 <130> FILE REFERENCE: PCT25390
 7 <140> CURRENT APPLICATION NUMBER: US/10/520,401
 8 <141> CURRENT FILING DATE: 2005-01-06
 9 <150> PRIOR APPLICATION NUMBER: IT RM2002A000365
10 <151> PRIOR FILING DATE: 2002-07-08
11 <160> NUMBER OF SEQ ID NOS: 9
12 <170> SOFTWARE: PatentIn version 3.2
14 <210> SEQ ID NO: 1
15 <211> LENGTH: 1265
16 <212> TYPE: DNA
17 <213> ORGANISM: Xenopus laevis
18 <220> FEATURE:
19 <221> NAME/KEY: misc_feature
20 <222> LOCATION: (1)..(38)
21 <223> OTHER INFORMATION: 5 untranslated DNA region
22 <220> FEATURE:
23 <221> NAME/KEY: exon
24 <222> LOCATION: (39)..(915)
25 <220> FEATURE:
26 <221> NAME/KEY: terminator
27 <222> LOCATION: (916)..(918)
28 <220> FEATURE:
29 <221> NAME/KEY: misc_feature
30 <222> LOCATION: (916)..(1265)
31 <223> OTHER INFORMATION: 3 untranslated DNA region
32 <400> SEQUENCE: 1
33       attggggaac tgggagcaga gagtgacggg caggagcc atg gcg agt aac agg ggg      56
34                                                  Met Ala Ser Asn Arg Gly
35                                                  1                   5
36       cag ctg aac cat gaa ctc tcc aag ctg ttt aat gag ctg tgg gac gca      104
37       Gln Leu Asn His Glu Leu Ser Lys Leu Phe Asn Glu Leu Trp Asp Ala
38                   10              15                  20
39       gat cag aac cgg atg aag tcc ggg aag gat tat cgg atc tcc ttg cag      152
40       Asp Gln Asn Arg Met Lys Ser Gly Lys Asp Tyr Arg Ile Ser Leu Gln
41               25                  30                  35
42       ggt aaa gca ggg tac gta ccc gcc ggt tcc aac cag gcc agg gac agc      200
43       Gly Lys Ala Gly Tyr Val Pro Ala Gly Ser Asn Gln Ala Arg Asp Ser
44           40                  45                  50
```

```
45   gcc tcg ttc ccg ctc ttc cag ttc gtc gat gag gag aag ctg aag agc   248
46   Ala Ser Phe Pro Leu Phe Gln Phe Val Asp Glu Glu Lys Leu Lys Ser
47   55                      60                  65                  70
48   agg aag acg ttt gca acc ttc att tcc ctg ctg gac aat tat gag atg   296
49   Arg Lys Thr Phe Ala Thr Phe Ile Ser Leu Leu Asp Asn Tyr Glu Met
50                       75                  80                  85
51   gac acg ggg gtg gcc gag gtt gtg act ccg gag gaa atc gct gaa aac   344
52   Asp Thr Gly Val Ala Glu Val Val Thr Pro Glu Glu Ile Ala Glu Asn
53                       90                  95                  100
54   aac aac ttc ctg gac gcc att ctg gaa acc aaa gtg atg aag atg gca   392
55   Asn Asn Phe Leu Asp Ala Ile Leu Glu Thr Lys Val Met Lys Met Ala
56               105                 110                 115
57   cat gac tac ctg gtg agg aag aac caa gcc aaa ccc acc cgg aat gac   440
58   His Asp Tyr Leu Val Arg Lys Asn Gln Ala Lys Pro Thr Arg Asn Asp
59           120                 125                 130
60   ttc aag gtc caa ctg tac aac atc tgg ttc cag ctg tac tca cgg gcc   488
61   Phe Lys Val Gln Leu Tyr Asn Ile Trp Phe Gln Leu Tyr Ser Arg Ala
62   135                 140                 145                 150
63   cca ggg agc aga ccc gat tcg tgc ggc ttt gag cac gtg ttt gtg gga   536
64   Pro Gly Ser Arg Pro Asp Ser Cys Gly Phe Glu His Val Phe Val Gly
65                       155                 160                 165
66   gaa tcg aag cga ggg cag gag atg atg ggg ctt cac aac tgg gtc cag   584
67   Glu Ser Lys Arg Gly Gln Glu Met Met Gly Leu His Asn Trp Val Gln
68                       170                 175                 180
69   ttt tac ctt cag gag aag agg aag aac atc gac tat aaa gga tac gtg   632
70   Phe Tyr Leu Gln Glu Lys Arg Lys Asn Ile Asp Tyr Lys Gly Tyr Val
71               185                 190                 195
72   gct cgg cag aac aag agt cgg ccg gat gaa gat gat cag gtg ttg aac   680
73   Ala Arg Gln Asn Lys Ser Arg Pro Asp Glu Asp Asp Gln Val Leu Asn
74           200                 205                 210
75   ctg cag ttc aat tgg aag gag atg gtg aaa ccc gtc ggc agc agc ttc   728
76   Leu Gln Phe Asn Trp Lys Glu Met Val Lys Pro Val Gly Ser Ser Phe
77   215                 220                 225                 230
78   att ggc gtc agc ccg gaa ttc gaa ttc gcc ctt tac acc atc gtc ttc   776
79   Ile Gly Val Ser Pro Glu Phe Glu Phe Ala Leu Tyr Thr Ile Val Phe
80                       235                 240                 245
81   ctc gcg tct cag gag aag atg agc cga gaa gtc gtt cgg ctg gaa gaa   824
82   Leu Ala Ser Gln Glu Lys Met Ser Arg Glu Val Val Arg Leu Glu Glu
83               250                 255                 260
84   tac gaa ctg cag atc gtc gtc aat cgc cac ggc cgt tat ata ggg acc   872
85   Tyr Glu Leu Gln Ile Val Val Asn Arg His Gly Arg Tyr Ile Gly Thr
86               265                 270                 275
87   gcc tac ccc gtc ctc ctg agc acc aat aac ccg gat ctg tac t         915
88   Ala Tyr Pro Val Leu Leu Ser Thr Asn Asn Pro Asp Leu Tyr
89           280                 285                 290
90   gagggggcgg ggctagagat cacagccggt tcccacggtt tgggtgcatt tactaacaaa   975
91   actgcaccaa tgcaacaatg caagcagata atggggggcag gtccatatcc ctctgctttc  1035
92   cctagcgtgt gtggggcaca ttaaccctat aactgtcact cactgcacca gacccattat  1095
93   ttaaccccac aagggacatc aagccagtgc cttgttatga gagagcgcag ccggggcttc  1155
```

```
94        tctactgtga aacttctgta ttgtatagag tttacttggt ttcttcctcc agacaatttc    1215
95        actttttttt tgctttgcct ttaaccatta aaagtccatg acatttctgt                1265
97  <210> SEQ ID NO: 2
98  <211> LENGTH: 292
99  <212> TYPE: PRT.
100 <213> ORGANISM: Xenopus laevis
101 <400> SEQUENCE: 2
102       Met Ala Ser Asn Arg Gly Gln Leu Asn His Glu Leu Ser Lys Leu Phe
103       1               5                   10                  15
104       Asn Glu Leu Trp Asp Ala Asp Gln Asn Arg Met Lys Ser Gly Lys Asp
105                       20                  25                  30
106       Tyr Arg Ile Ser Leu Gln Gly Lys Ala Gly Tyr Val Pro Ala Gly Ser
107               35                  40                  45
108       Asn Gln Ala Arg Asp Ser Ala Ser Phe Pro Leu Phe Gln Phe Val Asp
109           50                  55                  60
110       Glu Glu Lys Leu Lys Ser Arg Lys Thr Phe Ala Thr Phe Ile Ser Leu
111       65                  70                  75                  80
112       Leu Asp Asn Tyr Glu Met Asp Thr Gly Val Ala Glu Val Val Thr Pro
113                       85                  90                  95
114       Glu Glu Ile Ala Glu Asn Asn Asn Phe Leu Asp Ala Ile Leu Glu Thr
115                       100                 105                 110
116       Lys Val Met Lys Met Ala His Asp Tyr Leu Val Arg Lys Asn Gln Ala
117                   115                 120                 125
118       Lys Pro Thr Arg Asn Asp Phe Lys Val Gln Leu Tyr Asn Ile Trp Phe
119           130                 135                 140
120       Gln Leu Tyr Ser Arg Ala Pro Gly Ser Arg Pro Asp Ser Cys Gly Phe
121       145                 150                 155                 160
122       Arg His Val Phe Val Gly Glu Ser Lys Arg Gly Gln Glu Met Met Gly
123                   165                 170                 175
124       Leu His Asn Trp Val Gln Phe Tyr Leu Gln Glu Lys Arg Lys Asn Ile
125                   180                 185                 190
126       Asp Tyr Lys Gly Tyr Val Ala Arg Gln Asn Lys Ser Arg Pro Asp Glu
127           195                 200                 205
128       Asp Asp Gln Val Leu Asn Leu Gln Phe Asn Trp Lys Glu Met Val Lys
129           210                 215                 220
130       Pro Val Gly Ser Ser Phe Ile Gly Val Ser Pro Glu Phe Glu Phe Ala
131       225                 230                 235                 240
132       Leu Tyr Thr Ile Val Phe Leu Ala Ser Gln Glu Lys Met Ser Arg Glu
133                   245                 250                 255
134       Val Val Arg Leu Glu Glu Tyr Glu Leu Gln Ile Val Val Asn Arg His
135                   260                 265                 270
136       Gly Arg Tyr Ile Gly Thr Ala Tyr Pro Val Leu Leu Ser Thr Asn Asn
137           275                 280                 285
138       Pro Asp Leu Tyr
139           290
141 <210> SEQ ID NO: 3
142 <211> LENGTH: 20
143 <212> TYPE: RNA
144 <213> ORGANISM: Artificial Sequence
```

```
145 <220> FEATURE:
146 <223> OTHER INFORMATION: Synthetic oligoribonucleotide
147 <220> FEATURE:
148 <221> NAME/KEY: misc_RNA
149 <222> LOCATION: (1)..(20)
150 <223> OTHER INFORMATION: Synthetic oligoribonucleotide which includes U16 upstream
151       cleavage site
152 <400> SEQUENCE: 3
153       ggaaacguau ccuuugggag                                              20
155 <210> SEQ ID NO: 4
156 <211> LENGTH: 20
157 <212> TYPE: DNA
158 <213> ORGANISM: Artificial Sequence
159 <220> FEATURE:
160 <223> OTHER INFORMATION: Shyntetic oligonucleotide
161 <220> FEATURE:
162 <221> NAME/KEY: misc_RNA
163 <222> LOCATION: (1)..(20)
164 <223> OTHER INFORMATION: Mutant of SEQ ID 3
165 <220> FEATURE:
166 <221> NAME/KEY: variation
167 <222> LOCATION: (20)..(20)
168 <223> OTHER INFORMATION: "U" is replaced by "T"
169 <400> SEQUENCE: 4
170       ggaaacguau ccuugggagt                                              20
172 <210> SEQ ID NO: 5
173 <211> LENGTH: 20
174 <212> TYPE: RNA
175 <213> ORGANISM: Artificial Sequence
176 <220> FEATURE:
177 <223> OTHER INFORMATION: Synthetic oligoribonucleotide
178 <220> FEATURE:
179 <221> NAME/KEY: misc_RNA
180 <222> LOCATION: (1)..(20)
181 <223> OTHER INFORMATION: Mutant of SEQ ID 3
182 <220> FEATURE:
183 <221> NAME/KEY: variation
184 <222> LOCATION: (14)..(14)
185 <223> OTHER INFORMATION: "U" is replaced by "C"
186 <400> SEQUENCE: 5
187       ggaaacguau ccucugggag                                              20
189 <210> SEQ ID NO: 6
190 <211> LENGTH: 20
191 <212> TYPE: RNA
192 <213> ORGANISM: Artificial Sequence
193 <220> FEATURE:
194 <223> OTHER INFORMATION: Synthetic oligoribonucleotide
195 <220> FEATURE:
196 <221> NAME/KEY: misc_RNA
```

RAW SEQUENCE LISTING   DATE: 01/19/2005
PATENT APPLICATION: US/10/520,401   TIME: 20:32:42

Input Set : **N:\Crf4\01132005\J520401.raw**
Output Set: **N:\CRF4\01192005\J520401.raw**

```
197 <222> LOCATION: (1)..(20)
198 <223> OTHER INFORMATION: Mutant of SEQ ID 3
199 <220> FEATURE:
200 <221> NAME/KEY: variation
201 <222> LOCATION: (14)..(14)
202 <223> OTHER INFORMATION: "U" is replaced by "G"
203 <400> SEQUENCE: 6
204       ggaaacguau ccugugggag                                      20
206 <210> SEQ ID NO: 7
207 <211> LENGTH: 27
208 <212> TYPE: DNA
209 <213> ORGANISM: Artificial Sequence
210 <220> FEATURE:
211 <223> OTHER INFORMATION: Synthetic oligonucleotide
212 <220> FEATURE:
213 <221> NAME/KEY: misc_feature
214 <222> LOCATION: (1)..(27)
215 <223> OTHER INFORMATION: Primer
216 <400> SEQUENCE: 7
217       aagcttcttc atggcggctc ggccaat                              27
219 <210> SEQ ID NO: 8
220 <211> LENGTH: 20
221 <212> TYPE: DNA
222 <213> ORGANISM: Artificial Sequence
223 <220> FEATURE:
224 <223> OTHER INFORMATION: MAHs: synthetic degenerate oligonucleotide
225 <220> FEATURE:
226 <221> NAME/KEY: misc_feature
227 <222> LOCATION: (1)..(20)
228 <223> OTHER INFORMATION: PCR primer element
229 <220> FEATURE:
230 <221> NAME/KEY: modified_base
231 <222> LOCATION: (6)..(6)
232 <223> OTHER INFORMATION: "n" : I
233 <220> FEATURE:
234 <221> NAME/KEY: modified_base
235 <222> LOCATION: (18)..(18)
236 <223> OTHER INFORMATION: "n" : I
237 <400> SEQUENCE: 8
W--> 238       atggcncayg aytayytngt                                 20
240 <210> SEQ ID NO: 9
241 <211> LENGTH: 20
242 <212> TYPE: DNA
243 <213> ORGANISM: Artificial Sequence
244 <220> FEATURE:
245 <223> OTHER INFORMATION: IGTa: synthetic degenerate oligonucleotide
246 <220> FEATURE:
247 <221> NAME/KEY: misc_feature
248 <222> LOCATION: (1)..(20)
```

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 01/19/2005
PATENT APPLICATION: US/10/520,401   TIME: 20:32:43

Input Set : N:\Crf4\01132005\J520401.raw
Output Set: N:\CRF4\01192005\J520401.raw

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:8; N Pos. 6,18
Seq#:9; N Pos. 3,9,12,15,18

Invalid Line Length:
The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:1; Line(s) 3
Seq#:3; Line(s) 150

VERIFICATION SUMMARY   DATE: 01/19/2005
PATENT APPLICATION: US/10/520,401   TIME: 20:32:43

Input Set : N:\Crf4\01132005\J520401.raw
Output Set: N:\CRF4\01192005\J520401.raw

L:238 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:0
L:271 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9 after pos.:0